

Eine Studie zu den Merkmalen der Aufgabenschwierigkeit am Beispiel eines Informatik-Schülerwettbewerbs Zweiter Teil: Empirische Aufgabenanalyse

Kirsten Schlüter

Didaktik der Informatik
Universität Erlangen-Nürnberg
Martensstraße 3, 91058 Erlangen
schlueter@informatik.uni-erlangen.de

Zusammenfassung: Als Fortsetzung einer Studie zu den Schwierigkeitsmerkmalen von Informatikaufgaben wird in diesem Artikel die empirische Aufgabenanalyse anhand der Bearbeitungsdaten eines Informatik-Schülerwettbewerbs für die Sekundarstufe dargestellt. Die Wettbewerbsaufgaben wurden vorangehend bereits nach den vermuteten Kriterien der Aufgabenschwierigkeit klassifiziert. Auf der Basis der Ergebnisse dieser Aufgabenklassifizierung (Teil 1) und der empirischen Analyse (Teil 2), die die Schwierigkeit und die Trennschärfe der Aufgaben liefert, werden noch folgend die relevanten Schwierigkeitsmerkmale ausgewählt und die Dimensionen des Kompetenzanspruchs untersucht (Teil 3).

1 Zielsetzung

Mit dem Ziel, den Kompetenzanspruch von Informatikaufgaben zu modellieren, wird eine dreiteilige Studie durchgeführt, der die Teilnahmedaten des Online-Schülerwettbewerbs Informatik-Biber zu Grunde liegen. Zur Operationalisierung der spezifischen Kompetenzanforderungen von Informatikaufgaben soll ein Kriterienschema entwickelt werden, das es ermöglicht, Informatikaufgaben fundiert bezüglich ihrer Schwierigkeit einzuschätzen und in ein Modell des Kompetenzanspruchs einzuordnen. Das Schema soll sich an bewährten Taxonomien orientieren, auf den Aufgabenbestand anwendbar und im Alltag, nicht nur in der didaktischen Forschung, handhabbar sein, damit eine direkte Evaluation der Lernvoraussetzungen und Unterrichtserfolge im Schulfach Informatik durchgeführt werden kann. In diesem Artikel wird ausschließlich über den zweiten Teil der Studie, die empirische Aufgabenanalyse, berichtet. Auf die bereits abgeschlossene Klassifizierung der Aufgaben nach den Kriterien und die geplanten statistischen Analysen der Effekte der Aufgabenmerkmale und der Dimensionalität der Aufgabenschwierigkeit wird nur Bezug genommen, soweit es zur Einbettung der hier beschriebenen Untersuchungen erforderlich ist. Abbildung 1 zeigt die Arbeitsschritte der Studie in einem Schaubild.

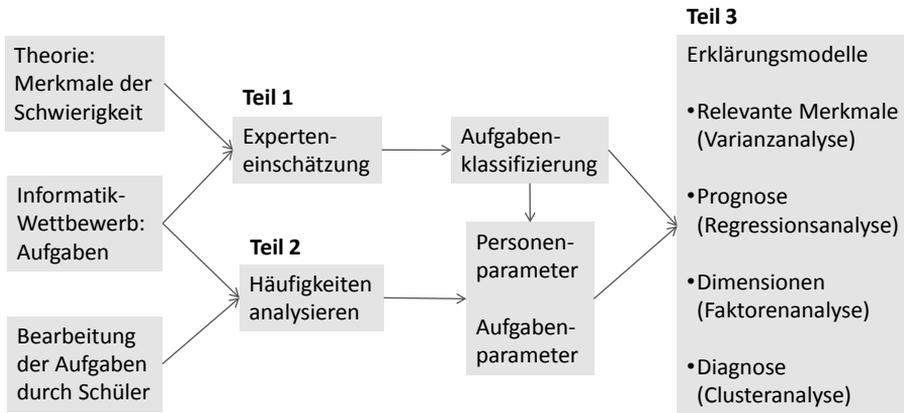


Abbildung 1: Methodik

Teil 1

Im bereits abgeschlossenen ersten Teil wurde aufgrund von Experteneinschätzungen die Klassifizierung der Wettbewerbsaufgaben nach charakterisierenden Merkmalen, die die Schwierigkeit und damit die spezifischen Kompetenzanforderungen bestimmen, vorgenommen. Als Schwierigkeitskriterien wurden zehn Merkmale begründet und untersucht: Inhaltsbereich, Erfahrungsweltnähe, Abstraktionsgrad, Komplexität, Formalisierungsgrad, Redundanz, Anforderungsbereich, Prozessbereich, kognitive Lernzielstufe und Art des Wissens [Sc09]¹. Zu den abgedruckten Aufgabenbeispielen sind in den Bildunterschriften auch die Klassifizierungsergebnisse aufgeführt (Abbildungen 2, 6, 9 und 10).

Teil 2

In diesem Artikel wird über den zweiten Teil der Studie berichtet, die empirische Aufgabenanalyse, bei der anhand der anonymisierten Bearbeitungsergebnisse der 20.000 Schüler und Schülerinnen der Sekundarstufe, die 2007 am Informatikwettbewerb teilgenommen haben, Schwierigkeitsparameter der Aufgaben und Personenparameter der Teilnehmer berechnet werden.

Teil 3

Erst im Abgleich der Teilstudienergebnisse, der Merkmalsklassifizierung (Teil 1) mit den Aufgabenparametern (Teil 2), kann der Zusammenhang der Merkmale mit der Aufgabenschwierigkeit überprüft werden. Dieses wird erst im dritten Teil durchgeführt, in dem Erklärungsmodelle betrachtet werden, die den Effekt der Schwierigkeitsmerkmale beschreiben. Erfolgversprechend sind Varianzanalysen zur Beurteilung der Relevanz der Merkmale, Faktorenanalysen zur Exploration der Dimensionalität des Kompetenzanspruchs, Regressionsanalysen zur Prognose der Aufgabenschwierigkeit, und schließlich Clusteranalysen zur Abgrenzung homogener Schülergruppen.

¹ Unter <http://ddi.informatik.uni-erlangen.de/research/kompetenzmodell/> stehen die versendeten Unterlagen sowie die Ergebnisse der Experteneinschätzung als Häufigkeitsverteilungen zur Verfügung.

2 Aufgabenanalyse: Datengrundlage

In der empirischen Aufgabenanalyse werden die Wettbewerbsaufgaben des Online-Schülerwettbewerbs Informatik-Biber untersucht. Der Biber ist das Einstiegsangebot des Bundeswettbewerbs Informatik, ein Projekt der Gesellschaft für Informatik (GI) und des Fraunhofer-Verbands IuK-Technologie, gefördert vom Bundesministerium für Bildung und Forschung. Der Wettbewerb wird seit 2006 jährlich ausgerichtet und ist an Schüler der Jahrgangsstufen 5 bis 13 adressiert. Die Aufgaben sind kurz und erfordern eine informatische Denkweise und Problembearbeitung, aber keine Vorkenntnisse in Informatik [Po09]². Abbildung 2 zeigt die Beispielaufgabe *Binärbaum*.

Der Wettbewerb wurde im Jahr 2007 in drei Altersgruppen durchgeführt: 5. bis 7., 8. bis 10., ab 11. Jahrgangsstufe. Jeder Altersgruppe lagen 15 der insgesamt 29 Aufgaben zur Bearbeitung vor. Wie in Tabelle 1 angedeutet, wurde für jeden der 20.000 Teilnehmer je Aufgabe erfasst, ob der Teilnehmer die Aufgabe korrekt (1) oder nicht korrekt gelöst (0) oder nicht bearbeitet hat (-). Die *Zeilensummen* entsprechen der Anzahl gelöster Aufgaben und geben einen Hinweis auf die Fähigkeit eines Teilnehmers, die Wettbewerbsaufgaben zu lösen. Die *Spaltensummen* entsprechen der Anzahl der Teilnehmer, die eine Aufgabe gelöst haben, und dienen als Ausgangswerte für die Schätzung der Aufgabenparameter Schwierigkeit und Trennschärfe aus den Wettbewerbsdaten.

Teilnehmer-ID	Jahrgangsstufe	Aufgabe 1	Aufgabe 2	...	Aufgabe 15	Zeilen-summe
1	5 bis 7	1	0	...	0	8
2	5 bis 7	-	1	...	0	8
3	8 bis 10	1	0	...	0	5
4	8 bis 10	1	-	...	1	13
5	ab 11	0	0	...	1	9
...
<i>Spaltensumme</i>		13555	6418	...	11765	

Tabelle 1: Datenauszug mit Zeilen- und Spaltensummen, Informatik-Biber 2007

Anzumerken ist, dass es sich bei dem Aufgabensatz nicht um einen psychometrischen Test handelt, der nach den Regeln der klassischen oder probabilistischen Testtheorie konstruiert wurde, wie etwa die Schulleistungsuntersuchungen der OECD PISA-Studie [OE06]. Vielmehr ist der Informatik-Biber 2007 ein informeller Leistungstest, der nur erlaubt, die Teilnehmer bezüglich ihrer Testergebnisse zu vergleichen. Sollen weitergehende Rückschlüsse auf Eigenschaften und Fähigkeiten der Teilnehmer gezogen werden, ist der Aufgabensatz auf Reliabilität und Validität zu überprüfen, sowie Durchführung und Auswertung des Wettbewerbs auf ihre Objektivität. Die Fragestellung dieser Studie betrifft aber zunächst nicht die Wettbewerbsteilnehmer, sondern die Aufgaben. Welche Merkmale sind es, die ihre Schwierigkeit beeinflussen und wie wirken sie zusammen? Die Beantwortung erfordert zum einen die aus dem ersten Teil bereits vorliegende Klassifizierung der Aufgaben nach den postulierten Schwierigkeitsmerkmalen [Sc09] und

² Die Webseite des Informatik-Bibers ist unter <http://www.informatik-biber.de> zu finden. Im Archiv stehen die Aufgaben der Wettbewerbe 2007 bis 2009 zur Verfügung.

zum anderen die hier beschriebene Analyse der Aufgaben im Hinblick auf Aufgabenparameter wie Schwierigkeit und Trennschärfe. Erst wenn die Teilergebnisse zusammengeführt und Effekte der Merkmale nachgewiesen wurden, richtet sich das Augenmerk auf die individuellen Wettbewerbsprofile der Teilnehmer.

Dies ist ein Beispiel für einen Binärbaum. Ein Binärbaum hat eine Wurzel (die ist oben – hier „A“), von der maximal zwei Äste abgehen. Am Ende eines Astes ist immer genau ein Knoten (hier „B“ bis „G“). Von jedem Knoten gehen wiederum maximal zwei Äste ab. Der Binärbaum rechts wird auch durch folgende Zeichenkette beschrieben:
 (A (B (C)) (D (E (F)) (G)))

Welcher der unten gezeigten Binärbäume wird durch folgende Zeichenkette beschrieben?
 (■(●)(▲(☾)(★)))

A)

B)

C)

D)

Abbildung 2: Aufgabe *Binärbaum*, Informatik-Biber 2007, Jahrgangsstufe 8 bis 10. Inhaltsbereich Information und Daten (Graphen), erfahrungsweltfern, konkret, mittel komplex, formal, nicht redundant, Anforderungsbereich Anwendung, Prozessbereich Darstellen und Interpretieren, Lernstufen Anwenden und Analysieren, Wissensart Konzepte.

3 Aufgabenanalyse: Kennzahlen

Unter der Schwierigkeit einer Aufgabe soll die empirische Schwierigkeit verstanden werden, das ist der Anteil der Personen, die die Aufgabe korrekt lösen. Unter der Trennschärfe soll der Grad verstanden werden, in dem eine Aufgabe zwischen den Testpersonen differenziert, das heißt, Personen mit hoher Leistung im Gesamttest lösen die Aufgabe deutlich öfter als Personen mit geringer Gesamtleistung. Während die Aufgabenschwierigkeit als Zielvariable in die statistischen Analysen eingeht, dient die Trennschärfe als Maß für die Aufgabengüte und als Kriterium, welche Aufgaben in die Interpretation der statistischen Ergebnisse einbezogen werden.

3.1 Aufgabenschwierigkeit

Der einfachste Schwierigkeitsindex ist der prozentuale Anteil der Personen, die eine Aufgabe lösen, $P = N_R / N * 100$ (N_R ist die Anzahl der Personen, die die betreffende Aufgabe gelöst haben – das ist die *Spaltensumme* in Tabelle 1. N ist die Anzahl der Personen, denen die Aufgabe vorlag). Hier wird zur Messung der Aufgabenschwierigkeit ein Schwierigkeitsindex mit Inangriffnahmekorrektur verwendet, $P_{IK} = N_R / N_B * 100$ (N_R wie oben, N_B ist die Anzahl der Personen, die die betreffende Aufgabe bearbeitet haben [Bü06]). Die Variante mit zusätzlicher Inangriffnahmekorrektur, die berücksich-

tigt wie viele Personen eine Aufgabe tatsächlich bearbeitet haben, wurde wegen der guten Passung zur Fragestellung gewählt. Systematische Vergleiche beider Varianten und weiterer Indizes ergaben nur geringfügige und regelmäßige Verschiebungen der Schwierigkeitswerte. In Tabelle 2 sind die Schwierigkeitswerte (Index) und deren Unterschiede zwischen den Jahrgangsstufen (Differenz) sowie die Trennschärfen (Koeffizient, siehe Kapitel 3.2) vollständig aufgelistet.

<i>Kennzahl:</i>		Index der Schwierigkeit (%)			Differenz (Δ Index)		Koeffizient der Trennschärfe (%)		
		5-7 (I)	8-10 (II)	11+ (III)	I, II	II, III	5-7 (I)	8-10 (II)	11+ (III)
1	Biber am Fluss	84,3					46,5		
2	Biber und Bisons		69,7	79,3		9,6		39,4	36,9
3	Bibers Geheimcode	39,9	56,1		16,2		49,7	48,8	
4	Biberzahlen		30,1	46,3		16,2		40,8	46,8
5	Binärbaum		59,6					46,0	
6	Computervirus	55,5	67,3	77,8	11,8	10,5	42,5	38,4	35,4
7	Dateisuche	47,5					43,8		
8	Dino-Ordnung	20,0					42,7		
9	Endlos-Schleife		47,7	66,1		18,4		39,6	46,8
10	Falschgeld		39,1	58,7		19,6		43,6	51,6
11	Fenster schließen	92,9					37,2		
12	Labyrinth			48,9					36,7
13	Link	77,4					47,4		
14	Links um!	68,8					51,4		
15	Morse-Code			25,7					39,1
16	Netzwerkabel		51,4	67,3		15,9		50,4	48,4
17	Platzwechsel	82,4					46,2		
18	POP und PUSH			51,7					48,0
19	Primärschlüssel		31,7	50,6		18,9		37,1	45,6
20	Private E-Mail		47,4					39,9	
21	Schnitzeljagd	35,3	42,4	57,4	7,1	15,0	39,0	47,5	46,9
22	Sicheres Passwort	69,5	68,9		-0,6		38,9	28,4	
23	Umparken	62,5	77,2		14,7		52,7	49,1	
24	Ungeschützter PC	21,7		24,7	Δ I, III: 3,0		26,7		27,1
25	Verschlüsselung		31,5	39,4		7,9		36,4	35,7
26	Verwandlung			43,4					51,5
27	Wertetausch			35,9					46,5
28	Wetter	23,9					34,9		
29	Zahlenreihe	33,5	46,9		13,4		47,5	47,6	
<i>Mittelwert:</i>		54,3	51,1	51,5	10,4	14,7	43,2	42,2	24,9
<i>Standardabweichung:</i>		24,5	15,0	16,7	6,2	4,3	6,9	6,1	7,2

Tabelle 2: Aufgabenkennzahlen. In den Jahrgangsstufen 5-7, 8-10, 11+ wurden jeweils 15 der insgesamt 29 Wettbewerbsaufgaben gestellt. Der Schwierigkeitsindex bezieht sich auf die Stufe, so dass es für mehrfach verwendete Aufgaben mehrere Schwierigkeitswerte gibt.

Zur übersichtlichen Darstellung in allen folgenden Diagrammen können die Aufgaben nach der Schwierigkeit in den Jahrgangsstufen geordnet werden. Die Werte reichen von 20% für *Dino-Ordnung* (d.h. 20% korrekte Lösungen, also vergleichsweise schwierig) bis 92,9% für *Fenster schließen* (d.h. 92,9% korrekte Lösungen, also eine sehr leichte Aufgabe). Werden die stufenübergreifend gestellten Aufgaben zur Verankerung genutzt, gelingt es bis auf wenige Konflikte sogar, eine gemeinsame Ordnung aller Aufgaben herzustellen. Abbildung 3 zeigt die geordneten Schwierigkeitswerte aller Aufgaben, mehrfach verwendete Aufgaben werden durch verbundene Punkte dargestellt.

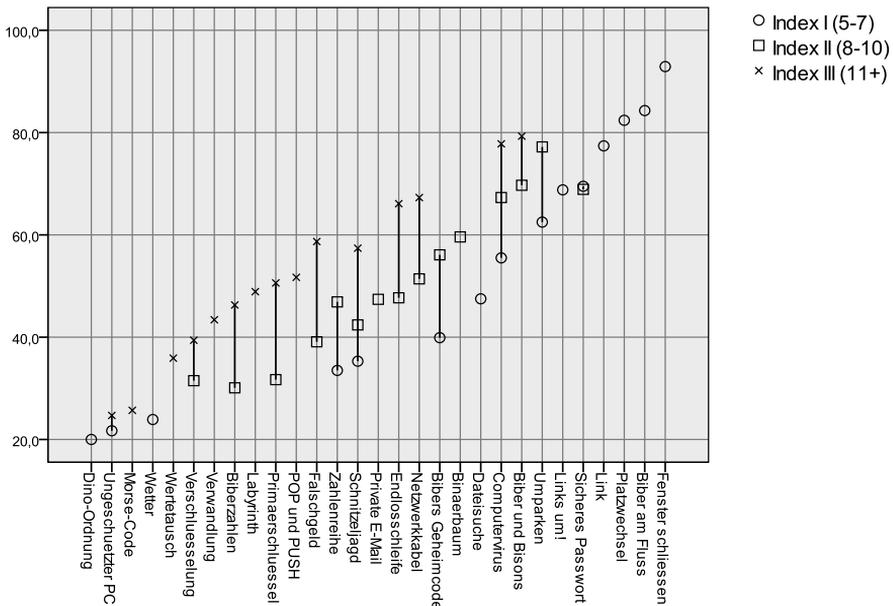


Abbildung 3: Anordnung der Aufgaben nach ihrer Schwierigkeit in der Stufe 5-7.

Konflikte ergeben sich nur vereinzelt bei Aufgaben, deren Schwierigkeitsunterschied zwischen den Jahrgangsstufen um mehr als eine Standardabweichung von der mittleren Differenz abweicht (Δ I/II, Δ II/III, Tabelle 2). So verletzt die Einordnung der Aufgabe *Schnitzeljagd* (Abbildung 10) gemäß Index I die Ordnung in den Stufen II und III. Solchen Aufgaben kommt besondere Aufmerksamkeit zu, weil sie Anhaltspunkte liefern, welche Merkmale die Aufgabenschwierigkeit in einer Altersgruppe beeinflussen.

3.2 Trennschärfe

Mögliche Maße für die Trennschärfe von Aufgaben sind der Korrelationskoeffizient zwischen Aufgabenergebnis und Gesamtergebnis (z.B. Punkt-biseriale Korrelation [BD05], Korrigierte Item-Skala-Interkorrelation [Bü06]) oder die Schwierigkeitsdifferenz zwischen den beiden Extremgruppen, den 25 % besten und den 25 % schlechtesten Testteilnehmern (z.B. Index of Discrimination [EF91]). Die Trennschärfe bemisst also, in welchem Grad die Aufgabe den Gesamtest repräsentiert. Eine trennscharfe Aufgabe wird häufiger von im Testsinne guten Schülern korrekt gelöst als von schlechten.

Vergleiche verschiedener Trennschärfe-Kennzahlen ergaben für die Wettbewerbsaufgaben lediglich Verhältnisunterschiede, keine Rangunterschiede. Für die Studie wurde die Korrelation zwischen dem Aufgabenergebnis und dem Personenparameter „Anzahl gelöster Aufgaben“ gewählt – das ist die *Zeilensumme* in Tabelle 1. Der Vorteil des Korrelationskoeffizienten besteht darin, dass die Variable, die als Testergebnis herangezogen wird, mit geringem Aufwand durch alternative Variablen ersetzt werden kann, etwa durch die Punktzahl, die im Wettbewerb vergeben wurde, und die auf theoretischen Schwierigkeitseinschätzungen der Aufgabensteller beruht. In Tabelle 2 sind neben den Schwierigkeitswerten die Trennschärfe-Koeffizienten aller Aufgaben aufgelistet.

Die Diagramme in Abbildung 4 vermitteln einen Überblick über die Schwierigkeit und Trennschärfe der Aufgaben in den drei Altersgruppen. Die Schwierigkeit der Aufgaben (schwarz durchgezogene Linie) streut in der ersten Gruppe am stärksten, von 20% bis 92,9%, in den anderen beiden Gruppen nur von 30,1% bis 77,2% beziehungsweise von 25,7% bis 79,3%, siehe auch Tabelle 2. Erwartungsgemäß weisen Aufgaben mittlerer Schwierigkeit höhere Trennschärfen auf als leichte und schwierige Aufgaben, die von allzu vielen Personen beziehungsweise nur von sehr wenigen Personen gelöst werden. Dies wird durch den Discrimination-Index (grau gestrichelte Linie) stärker betont als durch den Trennschärfe-Koeffizienten (schwarz gepunktete Linie). Korrelationstypisch gelten Werte im Bereich von 30% bis 50% (graue Bezugslinie) als mittlere, oberhalb davon als hohe Trennschärfen.

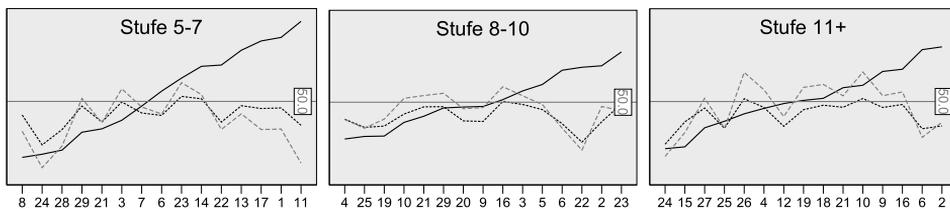


Abbildung 4: Schwierigkeitsindex (schwarz durchgezogene Linie), Trennschärfekoeffizient (schwarz gepunktet), Index of Discrimination (grau gestrichelt), 50%-Bezugslinie (grau).

3.3 Ergebnisse

Das Hauptergebnis der Aufgabenanalyse ist die empirische Schwierigkeit, Index I-III in Tabelle 2, die als Aufgabenparameter für die Berechnung der statistischen Erklärungsmodelle benötigt wird. Die Kennzahlen erschließen darüber hinaus bei geeigneter Darstellung das Aufgabenmaterial. Mit ihrer Hilfe lassen sich Aufgaben identifizieren, die zur Interpretation der statistischen Resultate beitragen. Das können die schwierigen Aufgaben sein, deren Merkmalsprofile eventuell Auffälligkeiten zeigen. Das können auch die Aufgaben sein, deren Schwierigkeitsentwicklung zwischen den Jahrgangsstufen Unregelmäßigkeiten aufweist. Und das können die Aufgaben mit hoher Trennschärfe sein, deren Anforderungsprofile die im Wettbewerb erfolgreichen Schüler besonders deutlich von den anderen Teilnehmern unterscheiden.

In den folgenden Streudiagrammen werden Aufgabenschwierigkeit und Trennschärfe der Wettbewerbsaufgaben miteinander verknüpft. Die Symbole enthalten die Information, in welchen der drei Altersgruppen eine Aufgabe verwendet wurde. Die „Verwendung“ zeigt lediglich an, ob die Datenlage einen stufenübergreifenden Vergleich ermöglicht. Sie ist nicht als Aufgabenmerkmal „Eignung“ zu verstehen, weil eine Aufgabe nicht unbedingt in jeder Altersgruppe verwendet wurde, für die sie geeignet erscheint (Abbildungen 5, 7 und 8).

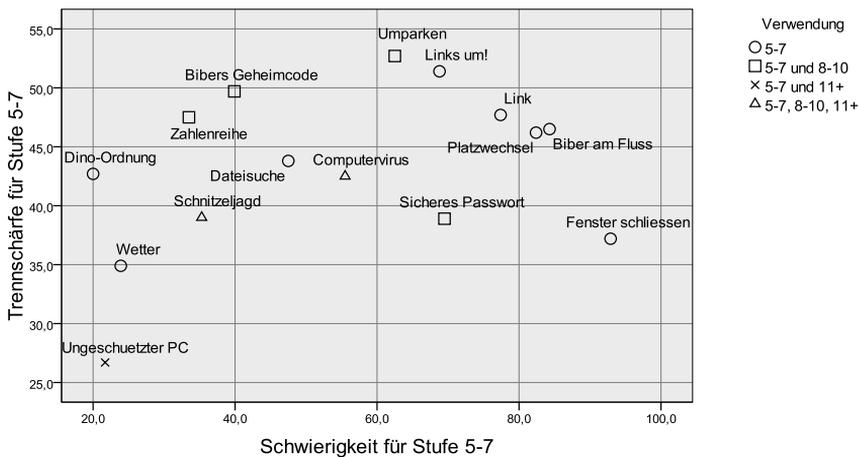


Abbildung 5: Aufgabenkennzahlen für die Gruppe I, Jahrgangsstufe 5-7

In der Altersgruppe I, 5. bis 7. Jahrgangsstufe, fällt auf dass beide Aufgaben zum Thema Logisches Schließen, *Dino-Ordnung* und *Wetter* (Abbildung 6) einen besonders hohen Schwierigkeitsgrad aufweisen. Es stellt sich die Frage, inwieweit diese Schwierigkeit durch den Inhaltsbereich, Schlussfolgern, erzeugt wird.

Angenommen, das Wetter folge der Regel:
 "Wenn an einem Tag die Sonne scheint, dann scheint auch am folgenden Tag die Sonne."



Wenn heute die Sonne scheint, was kannst du daraus folgern?

<p>A) Die Sonne schien bisher jeden Tag und wird auch jeden weiteren Tag scheinen.</p> <p>B) Gestern schien die Sonne.</p>	<p>C) Die Sonne wird nie wieder scheinen.</p> <p>D) Von heute an wird jeden Tag die Sonne scheinen.</p>
----------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------

Abbildung 6: Aufgabe *Wetter*, Informatik-Biber 2007, 5. bis 7. Jahrgangsstufe. Inhaltsbereich Automaten und Sprachen (Logisches Schließen als formales Kalkül), erfahrungsweltnah, konkret, einfach, informell, nicht redundant, Anforderungsbereich Anwendung, Prozessbereiche Begründen und Bewerten sowie Strukturieren und Vernetzen, Lernzielstufen Verstehen und Anwenden, Wissensart Konzepte.

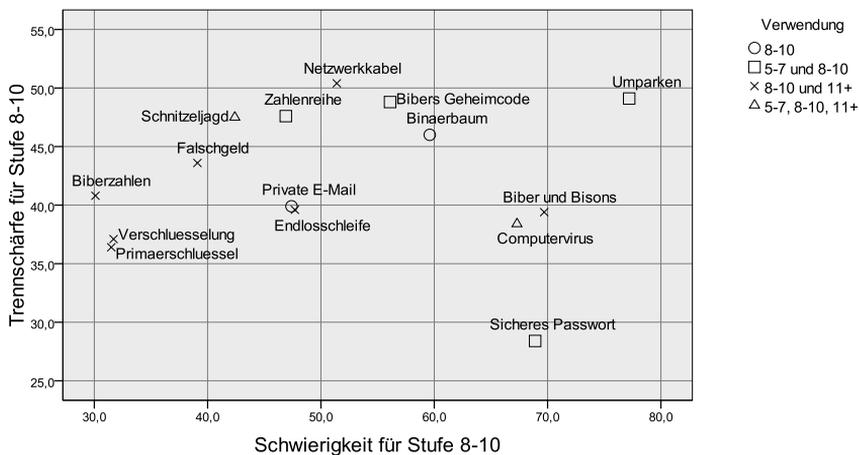


Abbildung 7: Aufgabenkennzahlen für die Gruppe II, Jahrgangsstufe 8-10

Ebenso besitzt die Aufgabe *Ungeschützter Computer* einen hohen Schwierigkeitsgrad, in der Gruppe I wie auch in der Gruppe III, ab 11. Jahrgangsstufe. Da Anwenderwissen erfragt wird, das in der 11. Jahrgangsstufe auch ohne schulischen Informatikunterricht vorhanden sein sollte, wird vermutet, dass die Schwierigkeit auf Gründe zurückgeht, die nicht Bestandteil dieser Untersuchung sind, etwa die missverständliche Formulierung der Antwortalternativen. Aufgaben wie diese werden mit entsprechendem Vorbehalt in den Analysen berücksichtigt.

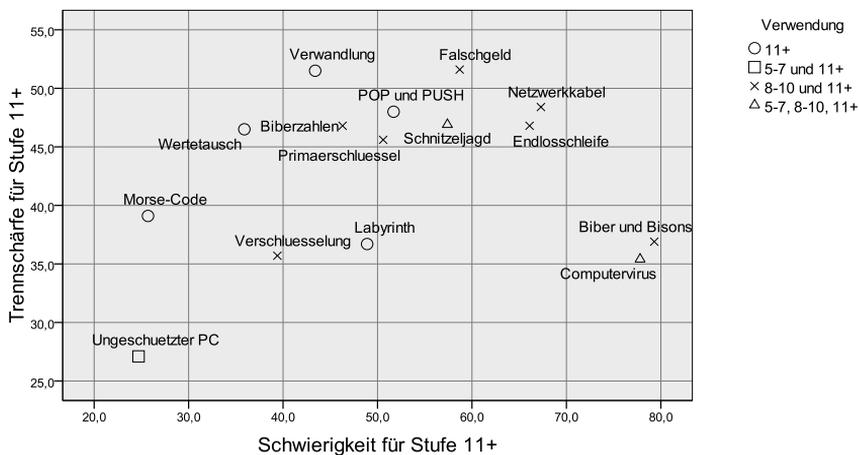


Abbildung 8: Aufgabenkennzahlen für die Gruppe III, Jahrgangsstufe 11+

Der Themenbereich Codierung stellt einige der schwierigsten Aufgaben in den Gruppen II und III, *Biberzahlen* und *Verschlüsselung* in der Jahrgangsstufe 8-10, *Morse-Code* in der Stufe 11+. Das gibt Anlass, die Aufgaben aus diesem Bereich auf spezifische Schwierigkeitsmerkmale zu untersuchen.

Umparken, mit dem Inhalt Datenstrukturen, und *Links um!*, aus dem Bereich Algorithmen (Abbildung 9), sind die trennschärfsten Aufgaben in der Gruppe I. *Umparken* wurde auch in der Altersgruppe II gestellt, wiederum mit hoher Trennschärfe, aber mit geringer Schwierigkeit. Ähnlich verhält es sich mit der Aufgabe *Netzwerkkabel*, Thema Graphen. Sie weist die höchste Trennschärfe in Stufe II auf und rangiert auch in Stufe III unter den Aufgaben mit guter Trennschärfe, bei geringerer Schwierigkeit. Die höchste Trennschärfe besitzt in Stufe III ebenfalls eine Graphen-Aufgabe, *Verwandlung*, neben einer Algorithmen-Aufgabe, *Falschgeld*.

Du hast einen Spielzeugroboter, der folgende Befehle auf Zuruf ausführen kann:

Befehl	Bedeutung
Vor!	Der Roboter fährt 10 cm nach vorne.
Rechts!	Der Roboter dreht sich nach rechts (um 90 Grad, also einen Viertelkreis).

Du möchtest nun den Roboter so bewegen, dass er am Ende um 90 Grad (einen Viertelkreis) nach links gedreht ist.

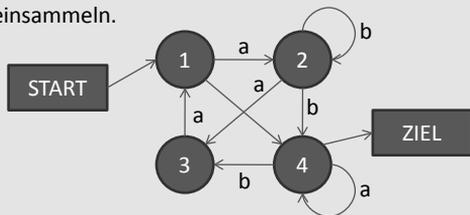
Mit welcher Befehlsfolge kannst du das erreichen?

- A) Vor! Vor! B) Rechts! Rechts! C) Rechts! Rechts! Rechts! D) Vor! Rechts! Vor!

Abbildung 9: Aufgabe *Links um!*, Informatik-Biber 2007, Jahrgangsstufe 5-7. Inhaltsbereich Algorithmen, mittel entfernt von der Erfahrungswelt, konkret, einfach, informell, nicht redundant, Anforderungsbereich Anwendung, Prozessbereich Modellieren und Implementieren, Lernzielstufe Anwenden. Die Wissensart konnte nicht übereinstimmend zugeordnet werden.

Die Aufgaben *Computervirus*, zum Thema Benutzerverhalten, und *Schnitzeljagd*, zum Thema Automaten (Abbildung 10), wurden in allen drei Altersgruppen eingesetzt und erwiesen sich als exemplarische Aufgaben mittlerer Schwierigkeit, die in der untersten Jahrgangsstufe am schwierigsten, in der oberen Jahrgangsstufe am leichtesten gelöst werden, und deren Trennschärfe im mittleren Schwierigkeitsbereich am deutlichsten ausgeprägt ist.

Auf seinem Weg vom Start zum Ziel folgt Florian den Pfeilen, beliebig lange. Jedes Mal, wenn er einen Pfeil entlang gegangen ist, sammelt er den zugehörigen Buchstaben ein. Und verlängert damit eine Kette der gesammelten Buchstaben. Bei einigen Pfeilen kann er keinen Buchstaben einsammeln.



Welche der folgenden Buchstabenketten kann Florian auf seinem Weg vom Start zum Ziel nicht einsammeln?

- A) abaabba C) abaaab
B) ba D) aab

Abbildung 10: Aufgabe *Schnitzeljagd*, Informatik-Biber 2007, Jahrgangsstufen 5-7, 8-10, ab 11. Inhaltsbereich Sprachen und Automaten, erfahrungsweltfern, konkret, mittel komplex, formal, nicht redundant, Anforderungsbereich Anwendung, Lernzielstufen Anwenden und Analysieren, Wissensart Prozeduren. Der Prozessbereich konnte nicht übereinstimmend zugeordnet werden.

3 Ausblick

Die 29 Wettbewerbsaufgaben wurden anhand der Bearbeitungsdaten von 20.000 Teilnehmern auf ihre Schwierigkeit und Trennschärfe untersucht. Diese Aufgabenparameter sowie Personenparameter, wie Anzahl und Merkmalsausprägung der gelösten Aufgaben, bilden gemeinsam mit der Aufgabenklassifizierung aus dem ersten Teil die nun vervollständigte Ausgangsbasis für den nächsten Schritt, die Erklärung des Kompetenzanspruchs durch charakteristische Aufgabenmerkmale. Dazu sind statistische Analysen zur Prüfung verschiedener Erklärungsmodelle geplant, zum Beispiel Regressionsanalysen, um die Aufgabenschwierigkeit aus den Kriterien vorherzusagen.

Einen rechnerischen Hinweis darauf, dass das durch die Aufgabenschwierigkeit erfasste Konstrukt nicht eindimensional ist, liefert die geringe Homogenität der Aufgaben: die Item-Interkorrelation beträgt in allen Altersgruppen weniger als 20%. Weitere Untersuchungen zur Dimensionalität sind erforderlich und werden im abschließenden dritten Teil der Studie durchgeführt. Es wird erwartet, mittels Faktorenanalysen eine mehrdimensionale Struktur der Aufgabenschwierigkeit und damit des Kompetenzanspruchs herausarbeiten zu können.

In die erste, visuelle Bestandsaufnahme der Aufgabenkennzahlen Schwierigkeit und Trennschärfe wurde bereits der Inhaltsbereich informell mit einbezogen. Das Merkmal Inhaltsbereich und die weiteren Aufgabenmerkmale Erfahrungsweltnähe, Abstraktionsgrad, Komplexität, Formalisierungsgrad, Redundanz, Anforderungsbereich, Prozessbereich, kognitive Lernzielstufe und Art des Wissens gehen systematisch in die Analyse und Interpretation der Ergebnisse ein.

In dem Fall dass die Faktorenanalyse nicht zu den gewünschten Faktoren führt, die als Dimensionen des Merkmalsraums interpretierbar sind, ist ein weiterer erfolgversprechender Ansatz zur Exploration des Schwierigkeitsgefüges, die Schülerinnen und Schüler mittels Clusteranalysen zu „ähnlichen“ Gruppen zusammenzufassen. Dabei bedeutet „ähnlich“, die Schüler haben Aufgaben gleichen Schwierigkeitsgrades bzw. Aufgaben mit den gleichen Schwierigkeitsmerkmalen gleich gut gelöst oder nicht gelöst. Von Interesse ist, ob typische Merkmalsprofile identifiziert werden können.

Weitere Indizien zur Interpretation können aus den Ergebnissen der Informatik-Biber-Wettbewerbe 2008 und 2009 gewonnen werden, die mittlerweile vorliegen. Einige Aufgaben aus 2007 wurden wieder verwendet, gegebenenfalls in anderen Altersgruppen, etwa die Aufgabe *Wetter*, die 2007 in Stufe 5-7 als besonders schwierig auffiel und 2008 in Stufe 11+ erneut gestellt wurde.

Danksagung

Vielen Dank an Dr. Wolfgang Pohl, Geschäftsführer des Bundeswettbewerbs Informatik, für die freundliche Erlaubnis, die anonymisierten Wettbewerbsdaten des Informatik-Bibers 2007 für die Studie zu verwenden.

Literaturverzeichnis

- [BD05] Bortz, J.; Döring, N.: Forschungsmethoden und Evaluation. Springer Medizin Verlag, Heidelberg, 2005.
- [Bü06] Bühner, M: Einführung in die Test- und Fragebogenkonstruktion. Pearson Studium, München, 2006.
- [EF91] Ebel, R. L.; Frisbie, D. A.: Essentials of Educational Measurement. Prentice Hall, Englewood Cliffs, N. J., 1991.
- [OE06] (OECD Hrsg.): Assessing Scientific, Reading and Mathematical Literacy. A Framework for PISA 2006. OECDpublishing, 2006; ISBN 978-92-64-02639-1.
- [Po09] Pohl, W. et al.: Informatik-Biber: Informatik-Einstieg und mehr. In (Koerber, B. Hrsg.): Zukunft braucht Herkunft. 25 Jahre „INFOS – Informatik und Schule“. 13. GI-Fachtagung Informatik und Schule 2009 in Berlin. Bonn, Köllen, 2009; S.38-49
- [Sc08] Schlüter, K.: Eine Studie zu den Merkmalen der Aufgabenschwierigkeit am Beispiel eines Informatik-Schülerwettbewerbs. Erster Teil: Aufgabenklassifizierung. In (Koerber, B. Hrsg.): Zukunft braucht Herkunft. 25 Jahre „INFOS – Informatik und Schule“. 13. GI-Fachtagung Informatik und Schule 2009 in Berlin. Bonn, Köllen, 2009; S.181-192.